30 | 服务器的管理和部署：工业界近几年有哪些发展趋势？

2020-02-03 庄振运

讲述：庄振运
时长 13:32　大小 10.86M

你好，我是庄振运。

说起服务器，你一定不陌生。那你知道 Facebook 的服务器是什么样的吗？要知道，Facebook 同时使用着很多不同的服务器。

在应对需要高速缓存的 Facebook 新闻、广告投放和搜索时，Facebook 使用的是有比较大内存和较强 CPU 的服务器。现在使用的每台服务器都有 256GB 的主内存和两个处理器的 CPU。

而在需要存储大量照片和视频的时候，Facebook 就选择了适用于数据和对象存储的服务器，这种服务器只有很少的内存，但是却有几百 TB 的硬盘存储空间。

今天，我们就从“服务器”入手，进入一个新的专题：容量管理工程。我们一起来看，要如何针对服务器设计、规划和部署的特点，开发出性能优越，能充分利用硬件资源的应用程序和服务。

如何设计一种新的服务器?

就像我前面用来举例的 Facebook 一样，大规模互联网公司的服务器，和我们家里以及办公室用的电脑可不一样，一般不是直接从市场上买的，而是自己设计的。

那么这些服务器是怎么设计出来的呢？其实服务器的设计和其他的硬件设计一样，也需要经过好几个阶段，你可以看一下它们的设计路线图（如下图所示）。

在最终进入大规模批量生产 MP 阶段之前，服务器的设计需要经历四个阶段，也就是：

Pre-EVT（Pre-Engineering Validation Test）：预先工程验证测试

EVT（Engineering Validation Test）：工程验证测试

DVT（Design Validation Test）：设计验证测试

PVT（Production Validation Test）：生产验证测试

这之后才是批量生产 MP 阶段。

每个阶段的目的和任务都有所区别，具体来讲，**Pre-EVT 阶段**进行的是比较初级的模块测试，且往往是和硬件供应商一起进行的。这一阶段通过基准测试，**得出基本的模块性能**，为下一步的、更具体的设计提供性能数据。所以，此阶段一般只需要少数几台服务器原型就可以。

第二个阶段是 **EVT**，这个阶段是服务器开发的初期设计验证，重点在**考虑服务器设计的完整度是否有遗漏**。这些测试，包括功能和安全规格测试。这时候的服务器是样品，问题可能较多，测试可能会做 N 次。但是这一阶段需要有一个完整的服务器，会运行更加具体的基准测试，并根据测试结果，来调整服务器设计。一般需要几十台服务器来进行测试。

DVT 阶段的测试，需要生产厂商把服务器运送到数据中心。到此时，几乎所有的设计已全部完成，重点是找出可能的设计问题，确保所有的设计都符合规格。此时产品基本定型，可以进行负载测试，是为了发现在生产环境中可能出现的问题。这个阶段的测试可能需要用上百台服务器。

PVT 阶段更加注重生产流程。服务器需要用类似真实的生产线来生产，并且运送到数据中心。这时需要进一步，用更加真实的应用程序来测试。这个阶段一般就要有百台以上甚至千台的服务器了。

MP 阶段，就是批量生产了。到这个阶段，服务器硬件准备就绪，而且需要使用真实生产流程，来进行大规模生产和应用测试。所有设计及生产问题，应该没有任何遗漏及错误，成为正式面市的产品。

服务器设计的机遇在哪?

仅仅知道如何设计一种新服务器还不够，你还需要知道整个业界如今的趋势，才能给出最有性价比的服务器设计。虽然业界会有比如全球 DRAM 短缺这样的短期波动，但是长期来看，也是有迹可循的，比如摩尔定律的放缓，越来越大的 HDD 硬盘，新的 SSD 技术等。

这些变化给我们的服务器设计带来了挑战，但同时也带来了机遇。CPU、内存、硬盘、网络、单处理器等几个方面的变化，不仅仅影响各公司数据中心未来服务器的发展方向，也对公司的软件和服务设计有重大而深远的影响。

CPU 的趋势

从 CPU 的趋势来看，简单来讲，摩尔定律正在失效，业界对 CPU 的性能增强，一般是通过添加更多内核来增大吞吐量。对于英特尔和其他 CPU 厂商（比如 ARM）来说，这意味着，每一代处理器将拥有更多的内核，但运行频率一般要低于当前一代的处理器。之所以有这些趋势，归根结底，是因为用于生产芯片的材料所面临的挑战，以及物理方面的基本限制，例如，光子和电子的特性等。

对软件来说，传统上，我们比较看重针对单线程的性能进行优化。鉴于这一硬件趋势的影响，这样的优化策略需要进行调整。无论处理器的体系结构（x86、ARM 等）如何，**CPU 单位价格的性能和每瓦性能的提升都在逐代地显著放缓**。

所以，我们的性能优化策略，应该着重服务的**水平扩展性**，就是通过多线程和多服务器来扩展服务的总体容量。

DRAM 内存的趋势

DRAM 内存的趋势如何呢？在全球市场上，DRAM 内存仍然相对稀缺，价格很高。自 2017 年以来，每 GB 的价格已翻了一番以上。与影响 CPU 一样，相同的基本问题和挑战，也影响着 DRAM 性能和容量的提高。所以，各公司的服务器设计，都开始转向使用尽量少的 DRAM 内存。

一个好消息是，新型技术比如 NVM（非易失性内存）等也在成熟；这就为许多应用，提供了更便宜的 DRAM 替代品。现在已经有越来越多的服务器设计，开始**考虑使用 NVM**。

如果你的互联网服务和程序使用内存较多，并且 SSD 这样的快速存储还是不能满足服务的要求，除了尽量减少内存使用外，建议你考虑采用 NVM。

硬盘的趋势

硬盘方面，硬盘的存储容量继续增加，但增速开始降低。这种趋势，是因为硬盘行业面临的材料挑战，以及物理方面的基本限制（例如控制磁场）。

虽然存储容量变大，但是硬盘提供的 IO 性能，比如每秒随机访问 IOPS 的性能并没有增加。为了利用这些更大容量的驱动器，我们最好**把硬件和软件设计一体化**。比如，你可以把热数据缓存在闪存或内存中，并将相对较冷的数据存储在旋转硬盘上。

网络的趋势

与服务器功耗相比，数据中心的网络流量增长得更快。具体来说，与计算和存储有关的网络功耗不断增长，所以需要使用更快、更多的交换机间的网络链接。

另一方面，新的技术比如硅光子技术，进一步降低了成本，也引发了新一轮的创新和数据中心的网络升级。比如有的数据中心，已经开始使用 400G、800G 甚至 1.6T 的高速网络。

但是你要记住，不管网络技术如何发展，还是需要尽量地减少跨数据中心和跨机架的网络流量，也就是**尽量让网络在本地消化掉**。这里的“本地”可以是服务器本身、本机架、本数据大厅、本数据中心等等。为什么呢？还记得我以前说过的“带宽超订”吗（参考第 18 讲），因为越往外走，网络的带宽总是越小。

如何才能尽量让网络流量本地消化呢？这就需要你考察公司内部各种服务之间的数据交互，尽量让有大量数据交换的服务在一起部署（比如部署在同一个机架内）。

单处理器的趋势

回到服务器的趋势上，有一个有趣的趋势就是：未来是单处理器的天下。

以前服务器采用多个处理器，是因为每个处理器上的内核数目有限。所以，如果我们希望一台服务器有更多的计算能力，只能采用更多的处理器。

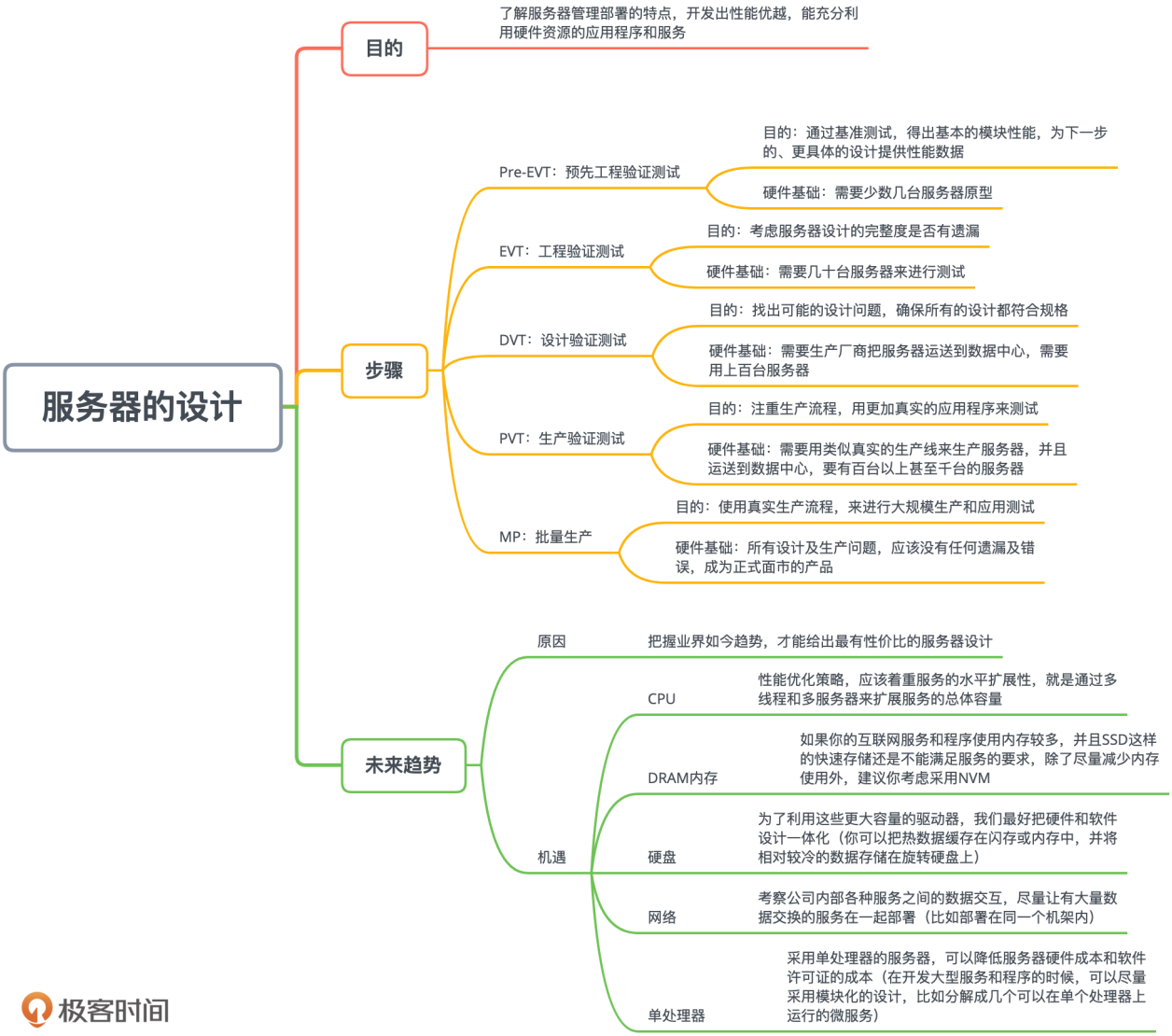
但是现在这种情况也正在改变，我们正进入高度集成的多核 CPU 时代，也就是一个处理器上面，有几十个内核越来越平常，一个处理器就已经有足够强大的计算能力。比如，随着每一代 x86 处理器的发布，都会增加更多的内核和功能，性能大大提高。

同时，服务器的内存密度也在增加。现在的单处理器服务器，能够置入的内存大小已经超过了以前服务器支持的内存密度。在很多生产环境中，内存容量和内存带宽，才是主要的性能瓶颈，而不是 CPU。因此，在双处理器环境中，CPU 使用率往往很低。

综合以上因素，采用单个处理器的服务器，可以降低服务器硬件成本和软件许可证的成本，让各种硬件资源得到更加充分的使用。具体来说，我们在开发大型服务和程序的时候，可以尽量采用模块化的设计，比如分解成几个可以在单个处理器上运行的微服务。

总结

我们这一讲，介绍了服务器设计的不同阶段，讨论了工业界这几年的发展趋势，包括CPU、内存、网络、磁盘等不同的服务器资源类型。



这些趋势，对于我们去把握下一代服务器硬件会有帮助。我们开发的互联网服务，总归是要运行在这些服务器上面，了解了这些趋势，我们才能开发出性能优越，能充分利用硬件资源的应用程序和服务。

清朝的赵翼有一首诗说：“李杜诗篇万口传，至今已觉不新鲜。江山代有才人出，各领风骚数百年。”其实服务器的发展，又何尝不是如此呢？服务器的更新换代速度，虽然可能没有软件那么快，但是也是几乎每年都有很大变化的，每种服务器，也是“各领风骚一两年”。

思考题

你们公司的服务器有几种？是不是也是按照资源的大小，比如内存和存储的大小来划分的呢？

你如果在公司已经工作几年了，有没有感受到硬件更新换代的特点？如果你在开发程序和互联网服务，想一想怎么设计你的程序和服务才能充分利用这些特点呢？

欢迎你在留言区分享自己的思考，与我和其他同学一起讨论，也欢迎你把文章分享给自己的朋友。

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇　29 | 如何彻底发挥SSD的潜力？

下一篇　31 | 规划部署数据中心要考虑哪些重要因素？

精选留言

写留言

由作者筛选后的优质留言将会公开显示，欢迎踊跃留言。